

# Bounds on the Voter Model in Dynamic Networks

Petra Berenbrink<sup>\*1</sup>, George Giakkoupis<sup>†2</sup>, Anne-Marie Kermarrec<sup>‡2</sup>, and  
Frederik Mallmann-Trenn<sup>§1,3</sup>

<sup>1</sup>Simon Fraser University, Burnaby, Canada

<sup>2</sup>INRIA, Rennes, France

<sup>3</sup>École Normale Supérieure, Paris, France

May 31, 2016

## Abstract

In the *voter model*, each node of a graph has an opinion, and in every round each node chooses independently a random neighbour and adopts its opinion. We are interested in the *consensus time*, which is the first point in time where all nodes have the same opinion. We consider dynamic graphs in which the edges are rewired in every round (by an adversary) giving rise to the graph sequence  $G_1, G_2, \dots$ , where we assume that  $G_i$  has conductance at least  $\phi_i$ . We assume that the degrees of nodes don't change over time as one can show that the consensus time can become super-exponential otherwise. In the case of a sequence of  $d$ -regular graphs, we obtain asymptotically tight results. Even for some static graphs, such as the cycle, our results improve the state of the art. Here we show that the expected number of rounds until all nodes have the same opinion is bounded by  $O(m/(d_{\min} \cdot \phi))$ , for any graph with  $m$  edges, conductance  $\phi$ , and degrees at least  $d_{\min}$ . In addition, we consider a *biased* dynamic voter model, where each opinion  $i$  is associated with a probability  $P_i$ , and when a node chooses a neighbour with that opinion, it adopts opinion  $i$  with probability  $P_i$  (otherwise the node keeps its current opinion). We show for any regular dynamic graph, that if there is an  $\epsilon > 0$  difference between the highest and second highest opinion probabilities, and at least  $\Omega(\log n)$  nodes have initially the opinion with the highest probability, then all nodes adopt w.h.p. that opinion. We obtain a bound on the convergence time, which becomes  $O(\log n/\phi)$  for static graphs.

## 1 Introduction

In this paper, we investigate the spread of opinions in a connected and undirected graph using the *voter model*. The standard voter model works in synchronous rounds and is defined as follows. At the beginning, every node has one opinion from the set  $\{0, \dots, n-1\}$ , and in every round, each node chooses one of its neighbours uniformly at random and adopts its opinion. In this model, one is usually interested in the *consensus time* and the *fixation probability*. The consensus time is the number of rounds it takes until all nodes have the same opinion. The fixation probability of opinion  $i$  is the probability that this opinion prevails, meaning that all other opinions vanish. This probability is known to be proportional to the sum of the degrees of the nodes starting with opinion  $i$  [15, 26].

The voter model is the dual of the *coalescing random walk model* which can be described as follows. Initially, there is a pebble on every node of the graph. In every round, every pebble chooses a neighbour uniformly at random and moves to that node. Whenever two or more pebbles meet at the same node, they are merged into a single pebble which continues performing a random walk. The process terminates when only one pebble remains. The time it takes until only one pebble remains is called *coalescing time*. It is known that the coalescing time for a graph  $G$  equals the consensus time of the voter model on  $G$  when initially each node has a distinct opinion [2, 21].

In this paper we consider the voter model and a *biased* variant where the opinions have different popularity. We express the consensus time as a function of the graph *conductance*  $\phi$ .

---

<sup>\*</sup>petra@sfu.ca

<sup>†</sup>george.giakkoupis@inria.fr

<sup>‡</sup>anne-marie.kermarrec@inria.fr

<sup>§</sup>fmallman@sfu.ca

We assume a dynamic graph model where the edges of the graph can be rewired by an adversary in every round, as long as the adversary respects the given degree sequence and the given conductance for all generated graphs. We show that consensus is reached with constant probability after  $\tau$  rounds, where  $\tau$  is the first round such that the sum of conductances up to round  $\tau$  is at least  $m/d_{\min}$ , where  $m$  is the number of edges. For static graphs the above bound simplifies to  $O(m/(d_{\min} \cdot \phi))$ , where  $d_{\min}$  is the minimum degree.

For the biased model we assume a *regular* dynamic graph  $G$ . Similar to [18, 21] the opinions have a *popularity*, which is expressed as a probability with which nodes adopt opinions. Again, every node chooses one of its neighbours uniformly at random, but this time it adopts the neighbour's opinion with a probability that equals the popularity of this opinion (otherwise the node keeps its current opinion). We assume that the popularity of the most popular opinion is 1, and every other opinion has a popularity of at most  $1 - \epsilon$  (for an arbitrarily small but constant  $\epsilon > 0$ ). We also assume that at least  $\Omega(\log n)$  nodes start with the most popular opinion. Then we show that the most popular opinion prevails w.h.p.<sup>1</sup> after  $\tau$  rounds, where  $\tau$  is the first round such that the sum of conductances up to round  $\tau$  is of order  $O(\log n)$ . For static graphs the above bound simplifies as follows: the most popular opinion prevails w.h.p. in  $O(\log n/\phi)$  rounds, if at least  $\Omega(\log n)$  nodes start with that opinion.

## 1.1 Related work

A sequential version of the voter model was introduced in [16] and can be described as follows. In every round, a single node is chosen uniformly at random and this node changes its opinion to that of a random neighbour. The authors of [16] study infinite grid graphs. This was generalised to arbitrary graphs in [11] where it is shown among other things that the probability for opinion  $i$  to prevail is proportional to the sum of the degrees of the nodes having opinion  $i$  at the beginning of the process.

The standard voter model was first analysed in [15]. The authors of [15] bound the expected coalescing time (and thus the expected consensus time) in terms of the expected meeting time  $t_{\text{meet}}$  of two random walks and show a bound of  $O(t_{\text{meet}} \cdot \log n) = O(n^3 \log n)$ . Note that the meeting time is an obvious lower bound on the coalescing time, and thus a lower bound on the consensus time when all nodes have distinct opinions initially. The authors of [6] provide an improved upper bound of  $O(\frac{1}{1-\lambda_2}(\log^4 n + \rho))$  on the expected coalescing time for any graph  $G$ , where  $\lambda_2$  is the second eigenvalue of the transition matrix of a random walk on  $G$ , and  $\rho = (\sum_{u \in V(G)} d(u))^2 / \sum_{u \in V(G)} d^2(u)$  is the ratio of the square of the sum of node degrees over the sum of the squared degrees. The value of  $\rho$  ranges from  $\Theta(1)$ , for the star graph, to  $\Theta(n)$ , for regular graphs.

The authors of [2, 21, 22] investigate coalescing random walks in a continuous setting where the movement of the pebbles are modelled by independent Poisson processes with a rate of 1. In [2], it is shown a lower bound of  $\Omega(m/d_{\max})$  and an upper bound of  $O(t_{\text{hit}} \cdot \log n)$  for the expected coalescing time. Here  $m$  is the number of edges in the graph,  $d_{\max}$  is the maximum degree, and  $t_{\text{hit}}$  is the (expected) hitting time. In [27], it is shown that the expected coalescing time is bounded by  $O(t_{\text{hit}})$ .

In [21] the authors consider the biased voter model in the continuous setting and two opinions. They show that for  $d$ -dimensional lattices the probability for the less popular opinion to prevail is exponentially small. In [18], it is shown that in this setting the expected consensus time is exponential for the line.

The authors of [7] consider a modification of the standard voter model with two opinions, which they call *two-sample voting*. In every round, each node chooses two of its neighbours randomly and adopts their opinion only if they both agree. For regular graphs and random regular graphs, it is shown that two-sample voting has a consensus time of  $O(\log n)$  if the initial imbalance between the nodes having the two opinions is large enough. There are several other works on the setting where every node contacts in every round two or more neighbours before adapting its opinion [1, 8, 9, 12].

There are several other models which are related to the voter model, most notably the *Moran process* and *rumor spreading* in the phone call model. In the case of the Moran process, a population resides on the vertices of a graph. The initial population consists of one mutant with fitness  $r$  and the rest of the nodes are non-mutants with fitness 1. In every round, a node is chosen at random with probability proportional to its fitness. This node then reproduces by placing a copy of itself on a randomly chosen neighbour, replacing the individual that was there. The main quantities of interest are the probability that the mutant occupies the whole graph (fixation) or vanishes (extinction), together with the time before either of the two states is reached (absorption time). There are several publications considering the fixation probabilities [10, 17, 23].

<sup>1</sup>An event happens *with high probability* (w.h.p.) if its probability is at least  $1 - 1/n$ .

Rumor spreading in the phone call model works as follows. Every node  $v$  opens a channel to a randomly chosen neighbour  $u$ . The channel can be used for transmissions in both directions. A transmission from  $v$  to  $u$  is called *push* transmission and a transmission from  $u$  to  $v$  is called *pull*. There is a vast amount of papers analysing rumor spreading on different graphs. The result that is most relevant to ours is that broadcasting of a message in the whole network is completed in  $O(\log n/\phi)$  rounds w.h.p, where  $\phi$  is the conductance (see Section 1.2 for a definition) of the network. In [14], the authors study rumor spreading in dynamic networks, where the edges in every round are distributed by an adaptive adversary. They show that broadcasting terminates w.h.p. in a round  $t$  if the sum of conductances up to round  $t$  is of order  $\log n$ . Here, the sequence of graphs  $G_1, G_2, \dots$  have the same vertex set of size  $n$ , but possibly distinct edge sets. The authors assume that the degrees and the conductance may change over time. We refer the reader to the next section for a discussion of the differences. Dynamic graphs have received ample attention in various areas [4, 19, 20, 28].

## 1.2 Model and New Results

In this paper we show results for the standard voter model and biased voter model in dynamic graphs. Our protocols work in synchronous steps. The consensus time  $T$  is defined at the first time step at which all nodes have the same opinion.

**Standard Voter Model.** Our first result concerns the standard voter model in dynamic graphs. Our protocol works as follows. In every synchronous time step every node chooses a neighbour u.a.r. and adopts its opinion with probability  $1/2$ .<sup>2</sup>

We assume that the dynamic graphs  $\mathcal{G} = G_1, G_2, \dots$  are generated by an adversary. We assume that each graph has  $n$  nodes and the nodes are numbered from 1 to  $n$ . The sequence of conductances  $\phi_1, \phi_2, \dots$  is given in advance, as well as a degree sequence  $d_1, d_2, \dots, d_n$ . The adversary is now allowed to create every graph  $G_i$  by redistributing the edges of the graph. The constraints are that each graph  $G_i$  has to have conductance  $\phi_i$  and node  $j$  has to have degree  $d_j$  (the degrees of the nodes do not change over time). Note that the sequence of the conductances is fixed and, hence, cannot be regarded as a random variable in the following. For the redistribution of the edges we assume that the adversary knows the distribution of all opinions during all previous rounds.

Note that our model for dynamic graphs is motivated by the model presented in [14]. They allow the adversary to determine the edge set at every round, without having to respect the node degrees and conductances.

We show (Observation 1) that, allowing the adversary to change the node degrees over time can result in super-exponential voting time. Since this changes the behaviour significantly, we assume that the degrees of nodes are fixed. Furthermore, in contrary to [14], we assume that (bounds on) the conductance of (the graph at any time step) are fixed/given beforehand. Whether one can obtain the same results, if the conductance of the graph is determined by an adaptive adversary remains an open question. The reason we consider an adversarial dynamic graph model is in order to understand how the voting time can be influenced in the worst-case. Another interesting model would be to assume that in every round the nodes are connected to random neighbours. One obstacle to such a model seems to be to guarantee that neighbours are chosen u.a.r. and the degrees of nodes do not change. For the case of regular random dynamic graphs our techniques easily carry over since the graph will have constant conductance w.h.p. in any such round since the graph is essentially a random regular graph in every round.

For the (adversarial) dynamic model we show the following result bounding the consensus time  $T$ .

**Theorem 1.1** (upper bound). *Consider the Standard Voter model and in the dynamic graph model. Assume  $\kappa \leq n$  opinions are arbitrarily distributed over the nodes of  $G_1$ . Let  $\phi_t$  be a lower bound on the conductance at time step  $t$ . Let  $b > 0$  be a suitable chosen constant. Then, with a probability of  $1/2$  we have that  $T \leq \min\{\tau, \tau'\}$ , where*

(i)  $\tau$  is the first round so that  $\sum_{t=1}^{\tau} \phi_t \geq b \cdot m/d_{\min}$ . (part 1)

(ii)  $\tau'$  is the first round so that  $\sum_{t=1}^{\tau'} \phi_t^2 \geq b \cdot n \log n$ . (part 2)

For static graphs ( $G_{i+1} = G_i$  for all  $i$ ), we have  $T \leq \min\{m/(d_{\min} \cdot \phi), n \log n/\phi^2\}$ .

For static  $d$ -regular graphs, where the graph doesn't change over time, the above bound becomes  $O(n/\phi)$ , which is tight when either  $\phi$  or  $d$  are constants (see Observation 2). Theorem 1.1 gives the first

---

<sup>2</sup>The factor of  $1/2$  ensures that the process converges on bipartite graphs.

tight bounds for cycles and circulant graphs  $C_n^k$  (node  $i$  is adjacent to the nodes  $i \pm 1, \dots, i \pm k \pmod n$ ) with degree  $2k$  ( $k$  constant). For these graphs the consensus time is  $\Theta(n^2)$ , which matches our upper bound from Theorem 1.1.<sup>3</sup> For a comparison with the results of [6] note that  $\phi^2 \leq 1 - \lambda_2 \leq 2\phi$ . In particular, for the cycle  $\phi = 1/n$  and  $1/(1 - \lambda_2) = \Theta(1/n^2)$ . Hence, for this graph, our bound is by a factor of  $n$  smaller. Note that, due to the duality between the voter model and coalescing random walks, the result also holds for the coalescing time. In contrast to [6, 7], the above result is shown using a potential function argument, whereas the authors of [6, 7] show their results for coalescing random walks and fixed graphs. The advantage of analysing the process directly is, that our techniques allow us to obtain the results for the dynamic setting.

The next result shows that the bound of Theorem 1.1 is asymptotically tight if the adversary is allowed to change the node degrees over time.

**Theorem 1.2** (lower bound). *Consider the Standard Voter model in the dynamic graph model. Assume that  $\kappa \leq n$  opinions are arbitrarily distributed over the nodes of  $G_1$ . Let  $\phi_t$  be an upper bound on the conductance at time step  $t$ . Let  $b > 0$  be a suitable constant and assume  $\tau''$  is the first round such that  $\sum_{t=1}^{\tau''} \phi_t \geq bn$ . Then, with a probability of at least  $1/2$ , there are still nodes with different opinions in  $G_{\tau''}$ .*

**Biased Voter Model** In the *biased* voter model we again assume that there are  $\kappa \leq n$  distinct opinions initially. For  $0 \leq i \leq \kappa - 1$ , opinion  $i$  has popularity  $\alpha_i$  and we assume that  $\alpha_0 = 1 > \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_{\kappa-1}$ . We call opinion 0 the *preferred opinion*. The process works as follows. In every round, every node chooses a neighbour uniformly at random and adopts its opinion  $i$  with probability  $\alpha_i$ .

We assume that the dynamic  $d$ -regular graphs  $\mathcal{G} = G_1, G_2, \dots$  are generated by an adversary. We assume that the sequence of  $\phi_t$  is given in advance, where  $\phi_t$  is a lower bound on the conductance of  $G_t$ . The adversary is now allowed to create the sequence of graphs by redistributing the edges of the graph in every step. The constraints are that each graph  $G_t$  has  $n$  nodes and has to have conductance at least  $\phi_t$ . Note that we assume that the sequence of the conductances is fixed and, hence, it is not a random variable in the following.

The following result shows that consensus is reached considerably faster in the biased voter model, as long as the bias  $1 - \alpha_1$  is bounded away from 0, and at least a logarithmic number of nodes have the preferred opinion initially.

**Theorem 1.3.** *Consider the Biased Voter model in the dynamic regular graph model. Assume  $\kappa \leq n$  opinions are arbitrarily distributed over the nodes of  $G_1$ . Let  $\phi_t$  be a lower bound on the conductance at time step  $t$ . Assume that  $\alpha_1 \leq 1 - \epsilon$ , for an arbitrary small constant  $\epsilon > 0$ . Assume the initial number of nodes with the preferred opinion is at least  $c \log n$ , for some constant  $c = c(\alpha_1)$ . Then the preferred opinion prevails w.h.p. in at most  $\tau'''$  steps, where  $\tau'''$  is the first round so that  $\sum_{t=1}^{\tau'''} \phi_t \geq b \log n$ , for some constant  $b$ . For static graphs ( $G_{i+1} = G_i$  for all  $i$ ), we have w.h.p.  $T = O(\log n / \phi)$ .*

The assumption on the initial size of the preferred opinion is crucial for the time bound  $T = O(\log n / \phi)$ , in the sense that there are instances where the expected consensus time is at least  $T = \Omega(n / \phi)$  if the size of the preferred opinion is small.<sup>4</sup>

The rumor spreading process can be viewed as an instance of the biased voter model with two opinions having popularity 1 and 0, respectively. However, the techniques used for the analysis of rumor spreading do not extend to the voter model. This is due to the fact that rumor spreading is a progressive process, where nodes can change their opinion only once, from “uninformed” to “informed”, whereas they can change their opinions over and over again in the case of the voter model. Note that the above bound is the same as the bound for rumor spreading of [13] (although the latter bound holds for general graphs, rather than just for regular ones). Hence, our above bound is tight for regular graphs with conductance  $\phi$ , since the rumor spreading lower bound of  $\Omega(\log n / \phi)$  is also a lower bound for biased voting in our model.

<sup>3</sup>The lower bound of  $\Omega(n^2)$  follows from the fact that two coalescing random walks starting on opposite sites of a cycle require in expectation time  $\Omega(n^2)$  to meet.

<sup>4</sup>Consider a 3-regular graph and  $n$  opinions where all other  $\alpha_1 = \alpha_2 = \dots = \alpha_{n-1} = 1/2$ . The preferred opinion vanishes with constant probability and the bound for the standard voter model of Observation 2 applies.

## 2 Analysis of the Voter Model

In this section we show the upper and lower bound for the standard voter model. We begin with some definitions. Let  $G = (V, E)$ . For a fixed set  $S \subseteq V$  we define  $\text{cut}(S, V \setminus S)$  to be the set of edges between the sets  $S \subseteq V$  and  $V \setminus S$  and let  $\lambda_u$  be the number of neighbours of  $u$  in  $V \setminus S$ . Let  $\text{vol}(S) = \sum_{u \in S} d_u$ . The *conductance* of  $G$  is defined as

$$\phi = \phi(G) = \min \left\{ \frac{\sum_{u \in U} \lambda_u}{\text{vol}(U)} : U \subset V \text{ with } 0 < \text{vol}(U) \leq m \right\}.$$

We note  $1/n^2 \leq \phi \leq 1$ . We denote by  $v_t^{(i)}$  the set of nodes that have opinion  $i$  after the first  $t$  rounds and  $t \geq 0$ . If we refer to the random variable we use  $V_t^{(i)}$  instead.

First we show Theorem 1.1 for  $\kappa = 2$  (two opinions), which we call 0 and 1 in the following. Then we generalise the result to an arbitrary number of opinions. We model the system with a Markov chain  $M_{t \geq 0} = (V_t^{(0)}, V_t^{(1)})_{t \geq 0}$ .

Let  $s_t$  denote the set having the smaller volume, i.e.,  $s_t = v_t^{(0)}$  if  $\text{vol}(v_t^{(0)}) \leq \text{vol}(v_t^{(1)})$ , and  $s_t = v_t^{(1)}$  otherwise. Note that we use  $s_t, v_t^{(0)}$  and  $v_t^{(1)}$  whenever the state at time  $t$  is fixed, and  $S_t, V_t^{(0)}$  and  $V_t^{(1)}$  for the corresponding random variables. For  $u \in v_t^{(0)}$ ,  $\lambda_{u,t}$  is the number of neighbours of  $u$  in  $V \setminus v_t^{(1)}$  and for  $u \in v_t^{(1)}$ ,  $\lambda_{u,t}$  is the number of neighbours of  $u$  in  $V \setminus v_t^{(0)}$ ;  $d_u$  is the degree of  $u$  (the degrees do not change over time).

To analyse the process we use a potential function. Simply using the volume of nodes sharing the same opinion as the potential function will not work. It is easy to calculate that the expected volume of nodes with a given opinion does not change in one step. Instead, we use a convex function on the number of nodes with the minority opinion. We define

$$\Psi(S_t) = \sqrt{\text{vol}(S_t)}.$$

In Lemma 2.1 we first calculate the one-step potential drop of  $\Psi(S_t)$ . Then we show that every opinion either prevails or vanishes once the sum of conductances is proportional to the volume of nodes having that opinion (see Lemma 2.2), which we use later to prove Part 1 and 2 of Theorem 1.1.

**Lemma 2.1.** *Assume  $s_t \neq \emptyset$  and  $\kappa = 2$ . Then*

$$\mathbf{E}[\Psi(S_{t+1}) \mid S_t = s_t] \leq \Psi(s_t) - \frac{\sum_{u \in V} \lambda_{u,t} \cdot d_u}{32 \cdot (\Psi(s_t))^3}.$$

*Proof.* W.l.o.g. we assume that opinion 0 is the minority opinion, i.e.  $0 < \text{vol}(V_t^{(0)}) \leq \text{vol}(V_t^{(1)})$ . To simplify the notation we omit the index  $t$  in this proof and write  $v^{(0)}$  instead of  $v_t^{(0)}$ ,  $v^{(1)}$  for  $V \setminus v_t^{(0)}$ , and  $\lambda_u$  instead of  $\lambda_{u,t}$ . Hence,  $s_t = v^{(0)}$  and  $\Psi(s_t) = \sqrt{\text{vol}(v^{(0)})}$ . Note that for  $t = 0$  we have  $\text{vol}(v^{(0)}) = \Psi(s_t)^2$ . Furthermore, we fix  $S_t = s_t$  in the following (and condition on it). We define  $m$  as the number of edges. Then we have

$$\begin{aligned} \mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] &= \mathbf{E}[\sqrt{\text{vol}(S_{t+1})} - \sqrt{\text{vol}(s_t)}] \\ &= \mathbf{E} \left[ \sqrt{\min \left\{ \text{vol}(V_{t+1}^{(0)}), m - \text{vol}(V_{t+1}^{(0)}) \right\}} - \sqrt{\text{vol}(s_t)} \right] \\ &\leq \mathbf{E} \left[ \sqrt{\text{vol}(V_{t+1}^{(0)})} - \sqrt{\text{vol}(v^{(0)})} \right] \end{aligned} \tag{1}$$

Now we define

$$X_u = \begin{cases} d_u & \text{w.p. } \frac{\lambda_u}{2 \cdot d_u} \text{ if } u \in v^{(1)} \\ -d_u & \text{w.p. } \frac{\lambda_u}{2 \cdot d_u} \text{ if } u \in v^{(0)} \\ 0 & \text{otherwise} \end{cases}$$

and  $\Delta = \sum_{u \in V} X_u$ . Note that we have  $\Delta = \text{vol}(V_{t+1}^{(0)}) - \text{vol}(v^{(0)})$  and

$$\mathbf{E} \left[ \sqrt{\text{vol}(V_{t+1}^{(0)})} - \sqrt{\text{vol}(v^{(0)})} \right] = \mathbf{E} \left[ \sqrt{\text{vol}(v^{(0)}) + \Delta} - \sqrt{\text{vol}(v^{(0)})} \right]$$

$$\begin{aligned}
&= \mathbf{E} \left[ \sqrt{\text{vol}(v^{(0)})} \left( \sqrt{1 + \frac{\Delta}{\text{vol}(v^{(0)})}} - 1 \right) \right] \\
&= \Psi(s_t) \cdot \mathbf{E}[\sqrt{1 + \Delta/\Psi(s_t)^2} - 1].
\end{aligned}$$

Unfortunately we cannot bound  $\Psi(s_t) \cdot \mathbf{E}[\sqrt{1 + \Delta/\Psi(s_t)^2} - 1]$  directly. Instead, we define a family of random variables which is closely related to  $X_u$ .

$$Y_u = \begin{cases} \lambda_u & \text{w.p. } \frac{1}{2} \quad \text{if } u \in v^{(1)} \\ -d_u & \text{w.p. } \frac{\lambda_u}{2 \cdot d_u} \quad \text{if } u \in v^{(0)} \\ 0 & \text{otherwise} \end{cases}$$

Similarly, we define  $\Delta' = \sum_{u \in V} Y(u)$ . Note that  $|E[Y_u]| = \lambda_u/2$  for both  $u \in v^{(1)}$  and  $u \in v^{(0)}$ . In Lemma A.1, we show that

$$\mathbf{E}[\sqrt{1 + \Delta/\Psi(s_t)^2}] \leq \mathbf{E}[\sqrt{1 + \Delta'/\Psi(s_t)^2}].$$

which results in  $\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] \leq \Psi(s_t) \cdot \mathbf{E}[\sqrt{1 + \Delta'/\Psi(s_t)^2} - 1]$  From the Taylor expansion  $\sqrt{1+x} \leq 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16}$ ,  $x \geq -1$  it follows that

$$\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] \leq \Psi(s_t) \cdot \mathbf{E} \left[ \frac{\Delta'}{2\Psi(s_t)^2} - \frac{(\Delta')^2}{8\Psi(s_t)^4} + \frac{(\Delta')^3}{16\Psi(s_t)^6} \right].$$

It remains to bound  $\mathbf{E}[\Delta']$ ,  $\mathbf{E}[(\Delta')^2]$ , and  $\mathbf{E}[(\Delta')^3]$ .

- $\mathbf{E}[\Delta']$ : We have  $\mathbf{E}[\Delta'] = \sum_{u \in V} E[Y_u] = \sum_{u \in v^{(1)}} \frac{\lambda_u}{2} - \sum_{u \in v^{(0)}} \frac{\lambda_u}{2} = 0$ , where the last equality holds since  $\sum_{u \in v^{(1)}} \lambda_u$  and  $\sum_{u \in v^{(0)}} \lambda_u$  both count the number of edges crossing the cut between  $v^{(0)}$  and  $v^{(1)}$ .
- $\mathbf{E}[(\Delta')^2]$ : since  $E[(Y_u)^2] = (\lambda_u)^2/2$  for  $u \in v^{(1)}$  and  $E[(Y_u)^2] = -d_u \cdot \lambda_u/2$  for  $u \in v^{(0)}$  we have

$$\begin{aligned}
\mathbf{E}[(\Delta')^2] &= \sum_{u \in V} \text{Var}[Y_u] + (\mathbf{E}[Y_u])^2 = \sum_{u \in V} \text{Var}[Y_u] + 0 = \sum_{u \in V} (\mathbf{E}[(Y_u)^2] - (\mathbf{E}[Y_u])^2) \\
&= \sum_{u \in v^{(0)}} (\mathbf{E}[(Y_u)^2] - (\mathbf{E}[Y_u])^2) + \sum_{u \in v^{(1)}} (\mathbf{E}[(Y_u)^2] - (\mathbf{E}[Y_u])^2) \\
&= \sum_{u \in v^{(0)}} \frac{\lambda_u d_u}{2} - \sum_{u \in v^{(0)}} \frac{\lambda_u^2}{4} + \sum_{u \in v^{(1)}} \frac{\lambda_u^2}{4} \geq \sum_{u \in v^{(0)}} \frac{\lambda_u d_u}{4}.
\end{aligned} \tag{2}$$

- $\mathbf{E}[\Delta'^3]$ : In Lemma A.2 we show that

$$\mathbf{E}[\Delta'^3] = \sum_{u \in V} (\mathbf{E}[(Y_u)^3] - 3\mathbf{E}[(Y_u)^2] \cdot \mathbf{E}[Y_u] + 2\mathbf{E}[Y_u]^3).$$

Note that  $E[(Y_u)^3] = \frac{1}{2}(\lambda_u)^3$  for  $u \in v^{(1)}$  and  $E[(Y_u)^3] = -\frac{1}{2}\lambda_u \cdot (d_u)^2$  for  $u \in v^{(0)}$ . Hence,

$$\begin{aligned}
\mathbf{E}[\Delta'^3] &= \sum_{u \in v^{(0)}} \left( -\frac{1}{2}\lambda_u \cdot (d_u)^2 + \frac{3}{4}(\lambda_u)^2 \cdot d_u - \frac{1}{4}\lambda_u^3 \right) \\
&\quad + \sum_{u \in v^{(1)}} \left( \frac{1}{2}(\lambda_u)^3 - \frac{3}{4}(\lambda_u)^3 + \frac{1}{4}(\lambda_u)^3 \right) \leq 0,
\end{aligned} \tag{3}$$

where the first sum is bounded by 0 because  $\lambda_u \leq d_u$ .

Combining all the above estimations we get

$$\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] \leq \Psi(s_t) \cdot \mathbf{E} \left[ \frac{\Delta'}{2\Psi(s_t)^2} - \frac{\Delta'^2}{8\Psi(s_t)^4} + \frac{\Delta'^3}{16\Psi(s_t)^6} \right] \leq -\frac{\sum_{u \in v^{(0)}} \lambda_u d_u}{32\Psi(s_t)^3}.$$

This completes the proof of Lemma 2.1.  $\square$

## 2.1 Part 1 of Theorem 1.1.

Using Lemma 2.1 we show that a given opinion either prevails or vanishes with constant probability as soon as the sum of  $\phi_t$  is proportional to the volume of the nodes having that opinion.

**Lemma 2.2.** Assume that  $s_{\hat{t}}$  is fixed for an arbitrary ( $\hat{t} \geq 0$ ) and  $\kappa = 2$ .

Let  $\tau^* = \min \left\{ t' : \sum_{i=\hat{t}}^{t'} \phi_i \geq 129 \cdot \text{vol}(s_{\hat{t}})/d_{\min} \right\}$ . Then  $\Pr(T \leq \tau^* + \hat{t}) \geq 1/2$ .

In particular, if the graph is static with conductance  $\phi$ , then  $\Pr(T \leq \frac{129 \cdot \text{vol}(s_{\hat{t}})}{\phi \cdot d_{\min}} + \hat{t}) \geq 1/2$ .

*Proof.* From the definition of  $\Psi(s_t)$  and  $\phi_t$  it follows for all  $t$  that  $\Psi(s_t)^2 = \sum_{u \in v^{(0)}} d_u = \text{vol}(v^{(0)})$  and  $\phi_t \leq \sum_{u \in v^{(0)}} \lambda_{u,t} / \text{vol}(v^{(0)})$ . Hence,  $\Psi(s_t)^2 \cdot \phi_t \cdot d_{\min} \leq \sum_{u \in v^{(0)}} \lambda_{u,t} \cdot d_u$ . Together with Lemma 2.1 we derive for  $s_t \neq \emptyset$

$$\mathbf{E}[\Psi(S_{t+1}) \mid S_t = s_t] \leq \Psi(s_t) - \frac{\sum_{u \in V} \lambda_{u,t} \cdot d_u}{32 \cdot (\Psi(s_t))^3} \leq \Psi(s_t) - \frac{d_{\min} \cdot \phi_t}{32 \cdot \Psi(s_t)}. \quad (4)$$

Recall that  $T = \min_t \{S_t = \emptyset\}$ . In the following we use the expression  $T > t$  to denote the event  $s_t \neq \emptyset$ . Using the law of total probability we get

$$\mathbf{E}[\Psi(S_{t+1}) \mid T > t] = \mathbf{E} \left[ \Psi(S_t) - \frac{d_{\min} \cdot \phi_t}{32 \cdot \Psi(S_t)} \mid T > t \right]$$

and using Jensen's inequality we get

$$\begin{aligned} \mathbf{E}[\Psi(S_{t+1}) \mid T > t] &= \mathbf{E}[\Psi(S_t) \mid T > t] - \mathbf{E} \left[ \frac{d_{\min} \cdot \phi_t}{32 \cdot \Psi(S_t)} \mid T > t \right] \\ &\leq \mathbf{E}[\Psi(S_t) \mid T > t] - \frac{d_{\min} \cdot \phi_t}{32 \cdot \mathbf{E}[\Psi(S_t) \mid T > t]}. \end{aligned}$$

Since  $\mathbf{E}[\Psi(S_t) \mid T \leq t] = 0$  we have

$$\begin{aligned} \mathbf{E}[\Psi(S_t)] &= \mathbf{E}[\Psi(S_t) \mid T > t] \cdot \Pr[T > t] + \mathbf{E}[\Psi(S_t) \mid T \leq t] \cdot \Pr[T \leq t] \\ &= \mathbf{E}[\Psi(S_t) \mid T > t] \cdot \Pr[T > t] + 0. \end{aligned}$$

Hence,

$$\frac{\mathbf{E}[\Psi(S_{t+1})]}{\Pr(T > t)} \leq \frac{\mathbf{E}[\Psi(S_t)]}{\Pr(T > t)} - \frac{d_{\min} \cdot \phi_t \cdot \Pr(T > t)}{32 \mathbf{E}[\Psi(S_t)]}$$

and

$$\mathbf{E}[\Psi(S_{t+1})] \leq \mathbf{E}[\Psi(S_t)] - \frac{d_{\min} \cdot \phi_t \cdot (\Pr(T > t))^2}{32 \mathbf{E}[\Psi(S_t)]}.$$

Let  $t^* = \min\{t : \Pr(T > t) < 1/2\}$ . In the following we use contradiction to show

$$t^* \leq \max\{t : \sum_{\hat{t} \leq i < t^*} \phi_i \leq 128 \cdot \text{vol}(s_{\hat{t}})/d_{\min}\}.$$

Assume the inequality is not satisfied. With  $t = t^* - 1$  we get

$$\mathbf{E}[\Psi(S_{t^*})] \leq \mathbf{E}[\Psi(S_{t^*-1})] - \frac{d_{\min} \cdot \phi_t \cdot (\Pr(T > t^* - 1))^2}{32 \mathbf{E}[\Psi(S_{t^*-1})]} \leq \mathbf{E}[\Psi(S_{t^*-1})] - \frac{d_{\min} \cdot \phi_{t^*} \cdot (1/4)}{32 \mathbf{E}[\Psi(S_{t^*-1})]}.$$

Applying this equation iteratively, we obtain

$$\mathbf{E}[\Psi(S_{t^*})] \leq \mathbf{E}[\Psi(S_{\hat{t}})] - \sum_{\hat{t} \leq i < t^*} \frac{d_{\min} \cdot \phi_i \cdot 1/4}{32 \mathbf{E}[\Psi(S_i)]} \leq \mathbf{E}[\Psi(S_{\hat{t}})] - \frac{d_{\min} \cdot \sum_{\hat{t} \leq i < t^*} \phi_i}{128 \mathbf{E}[\Psi(S_{\hat{t}})]}. \quad (5)$$

Using the definition of  $\mathbf{E}[\Psi(S_{\hat{t}})] = \sqrt{\text{vol}(s_{\hat{t}})}$  and the definition of  $t^*$  we get

$$\mathbf{E}[\Psi(S_{t^*})] < \sqrt{\text{vol}(s_{\hat{t}})} - \frac{d_{\min} \cdot 128 \cdot \text{vol}(s_{\hat{t}})}{128 \cdot d_{\min} \cdot \sqrt{\text{vol}(s_{\hat{t}})}} = \sqrt{\text{vol}(s_{\hat{t}})} - \frac{\text{vol}(s_{\hat{t}})}{\sqrt{\text{vol}(s_{\hat{t}})}} = 0.$$

This is a contradiction since  $\mathbf{E}[\Psi(S_{t^*})]$  is non-negative.

From the definition of  $t^*$ , we obtain  $\Pr(T > \tau^* + \hat{t}) < 1/2$ , completing the proof of Lemma 2.2.  $\square$

Now we are ready to show the first part of the theorem.

*Proof of Part 1 of Theorem 1.1.* We divide the  $\tau$  rounds into phases. Phase  $i$  starts at time  $\tau_i = \min\{t : \sum_{j=1}^t \phi_j \geq 2i\}$  for  $i \geq 0$  and ends at  $\tau_{i+1} - 1$ . Since  $\phi_j \leq 1$  for all  $j \geq 0$  we have  $\tau_0 < \tau_1 < \dots$  and  $\sum_{j=\tau_i}^{\tau_{i+1}-1} \phi_j \geq 1$  for  $i \geq 0$ . Let  $\ell_t$  be the number of distinct opinions at the beginning of phase  $t$ . Hence,  $\ell_0 = \kappa$ .

We show in Lemma 2.3 below that the expected number of phases before the number of opinions drops by a factor of  $5/6$  is bounded by  $6c \cdot \text{vol}(V)/(\ell_t \cdot d_{\min})$ . For  $i \geq 1$  let  $T_i$  be the number of phases needed so that the number of opinions drops to  $(5/6)^i \cdot \ell_0$ . Then only one opinion remains after  $\log_{6/5} \kappa$  many of these meta-phases. Then, for a suitably chosen constant  $b$ ,

$$\mathbf{E}[T] = \sum_{j=1}^{\log_{6/5} \kappa} \mathbf{E}[T_j] \leq \sum_{j=1}^{\log_{6/5} \kappa - 1} \frac{6c \cdot \text{vol}(V)}{\ell_j \cdot d_{\min}} \leq \sum_{j=1}^{\log_{6/5} \kappa} \frac{6c \text{vol}(V)}{(5/6)^j \cdot \ell_0 \cdot d_{\min}} = \frac{b \cdot m}{4 \cdot d_{\min}}.$$

By Markov inequality, consensus is reached w.p. at least  $1/2$  after  $b \cdot m/(2d_{\min})$  phases. By definition of  $\tau$  and the definition of the phases, we have that the number of phases up to time step  $\tau$  is at least  $b \cdot m/(2d_{\min})$ . Thus, consensus is reached w.p. at least  $1/2$  after  $\tau$  time steps, which finishes the proof.  $\square$

**Lemma 2.3.** *Fix a phase  $t$  and assume  $c = 129$  and  $\ell_t > 1$ . The expected number of phases before the number of opinions drops to  $5/6 \cdot \ell_t$  is bounded by  $6c \cdot \text{vol}(V)/(\ell_t \cdot d_{\min})$ .*

*Proof.* Consider a point when there are  $\ell'$  opinions left, with  $5/6 \cdot \ell < \ell' \leq \ell$ . Among those  $\ell'$  opinions, there are at least  $\ell' - \ell/3$  opinions  $i$  such that the volume of nodes with opinion  $i$  is at most  $3 \cdot \text{vol}(V)/\ell$ . Let  $S$  denote the set of these opinions and let  $Z_i$  be an indicator variable which is 1 if opinion  $i \in S$  vanished after  $s = 3c \cdot \text{vol}(V)/(\ell \cdot d_{\min})$  phases and  $Z_i = 0$  if it prevails. To estimate  $Z_i$  we consider the process where we have two opinions only. All nodes with opinion  $i$  retain their opinion and all other nodes have opinion 0. It is easy to see that in both processes the set of nodes with opinion  $i$  remains exactly the same. Hence, we can use Lemma 2.2 to show that with probability at least  $1/2$ , after  $s$  phases opinion  $i$  either vanishes or prevails. Hence,

$$\mathbf{E}[\sum_{j \in S} Z_j] = \sum_{j \in S} \mathbf{E}[Z_j] \geq |S|/2 \geq (\ell' - \ell/3)/2.$$

Using Markov's inequality we get that with probability  $1/2$  at least  $(\ell' - \ell/3)/4$  opinions vanish within  $s$  phases, and the number of opinions remaining is at most  $\ell' - (\ell' - \ell/3)/4 = 3/4 \cdot \ell' + \ell/12 \leq 5/6 \cdot \ell$ . The expected number of phases until  $5/6 \cdot \ell$  opinions can be bounded by  $\sum_{i=1}^{\infty} 2^{-i} \cdot s \leq 2s = \frac{6c \cdot \text{vol}(V)}{\ell \cdot d_{\min}}$ .  $\square$

## 2.2 Part 2 of Theorem 1.1

The following lemma is similar to Lemma 2.2 in the last section: We first bound the expected potential drop in round  $t + 1$ , i.e., we bound  $\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t]$ . This time however, we express the drop as a function which is linear in  $\Psi(s_t)$ . This allows us to bound the expected size of the potential at time  $\tau'$ , i.e.,  $\mathbf{E}[\Psi(S_{\tau'})]$ , directly. From the expected size of the potential at time  $\tau'$  we derive the desired bound on  $\Pr(T \leq \tau')$ .

**Lemma 2.4.** *Assume  $\kappa = 2$ . We have  $\Pr(T \leq \tau') \geq 1/n^2$ . In particular, if the graph is static with conductance  $\phi$ , then  $\Pr(T \leq \frac{96 \cdot n \log n}{\phi^2}) \geq 1 - 1/n^2$ .*

*Proof.* In the following we fix a point in time  $t$  and use  $\lambda_u$  instead of  $\lambda_{u,t}$ . From Lemma 2.1 and the observation  $\lambda_u \leq d_u$  we obtain for  $s_t \neq \emptyset$

$$\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] \leq -\frac{\sum_{u \in V} \lambda_u \cdot d_u}{32 \cdot (\Psi(s_t))^3} \leq -\frac{\sum_{v \in v^{(0)}} (\lambda_u)^2}{32(\Psi(s_t))^3}.$$

We have, by Cauchy-Schwarz inequality,  $(\sum_{v \in v^{(0)}} \lambda_u)^2 = (\sum_{v \in v^{(0)}} \lambda_u \cdot 1)^2 \leq \sum_{v \in v^{(0)}} (\lambda_u)^2 \cdot n$ . Hence,

$$\sum_{v \in V} (\lambda_u)^2 \geq \sum_{v \in v^{(0)}} (\lambda_u)^2 \geq \frac{(\sum_{v \in v^{(0)}} \lambda_u)^2}{n} \geq \frac{(\text{vol}(v^{(0)}))^2 \cdot (\phi_t)^2}{n} \geq \frac{\Psi(s_t)^4 \cdot (\phi_t)^2}{n},$$

where the third inequality follows by definition of  $\phi_t$ . Hence, for  $s_t \neq \emptyset$  we have

$$\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] \leq -\frac{\Psi(s_t) \cdot (\phi_t)^2}{32n}.$$



Note that  $\mathbf{E}[\Psi(S_{t+1}) \mid S_t = \emptyset] = 0 = (1 - \frac{(\phi_t)^2}{32n}) \cdot \mathbf{E}[\Psi(\emptyset)]$ . Hence, for all  $t \geq 1$  we get

$$\mathbf{E}[\Psi(S_{t+1})] = \mathbf{E}[\mathbf{E}[\Psi(S_{t+1}) \mid S_t = s_t]] \leq \left(1 - \frac{(\phi_t)^2}{32n}\right) \cdot \mathbf{E}[\Psi(s_t)].$$

Applying this recursively yields

$$\mathbf{E}[\Psi(S_{t+1})] \leq \Psi(S_0) \cdot \prod_{i=0}^t \left(1 - \frac{(\phi_i)^2}{32n}\right) \leq \Psi(S_0) \cdot \left(1 - \frac{1}{t+1} \sum_{i \leq t} \frac{(\phi_i)^2}{32n}\right)^{t+1} \leq \Psi(S_0) \cdot \exp\left(-\sum_{i \leq t} \frac{(\phi_i)^2}{32n}\right),$$

where the second inequality follows from the Inequality of arithmetic and geometric means.

By definition of  $\tau'$ , and from the observation  $\Psi(S_0) \leq n$  we get that  $\mathbf{E}[\Psi(S_{\tau'})] \leq n^{-2}$ .

We derive

$$n^{-2} \geq \mathbf{E}[\Psi(S_{\tau'})] \geq 0 \cdot \Pr(\Psi(S_{\tau'}) = 0) + 1 \cdot (1 - \Pr(\Psi(S_{\tau'}) = 0)), \quad (6)$$

where we used that  $\min\{\Psi(S) : S \subseteq V : S \neq \emptyset\} \geq 1$ . Solving (6) for  $\Pr(\Psi(S_{\tau'}) = 0)$  gives  $\Pr(\Psi(S_{\tau'}) = 0) \geq 1 - 1/n^2$ .

Since  $\kappa = 2$ , it follows that  $\Pr(T \leq \tau') = \Pr(\Psi(S_{\tau'}) = 0) \geq 1 - 1/n^2$ , which yields the claim.  $\square$

We now prove Part 2 of Theorem 1.1 which generalises to  $\kappa > 2$ .

*Proof of Part 2 of Theorem 1.1.* We define a parameterized version of the consensus time  $T$ . We define  $T(\kappa) = \min\{t : \Psi(S_t) = 0 : \text{the number of different opinions at time } t \text{ is } \kappa\}$  for  $\kappa \leq n$ . We want to show that  $\Pr(T(\kappa) \leq \tau') \geq 1 - 1/n$ . From Lemma 2.4 we have that, that  $\Pr(T(2) \leq \tau') \geq 1 - 1/n^2$ . We define the 0/1 random variable  $Z_i$  to be one if opinion  $i$  vanishes or is the only remaining opinion after  $\tau'$  rounds and  $Z_i = 0$  otherwise. We have that  $\Pr(Z_i = 1) \geq 1 - 1/n^2$  for all  $i \leq \kappa$ . We derive  $\Pr(T(\kappa) \leq \tau') = \Pr(\bigwedge_{i \leq \kappa} Z_i) \geq 1 - 1/n$ , by union bound. This yields the claim.  $\square$

## 2.3 Lower Bounds

In this section, we give the intuition behind the proof of Theorem 1.2 and state two additional observations. Recall that Theorem 1.2 shows that our bound for regular graphs is tight for the adaptive adversary, even for  $k = 2$ . The first observation shows that the expected consensus time can be super-exponential if the adversary is allowed to change the degree sequence. The second observation can be regarded as a (weaker) counter part of Theorem 1.2 showing a lower bound of  $\Omega(n/\phi)$  for static graphs, assuming that either  $d$  or  $\phi$  is constant.

We now give the intuition behind the proof of Theorem 1.2. The high level approach is as follows. For every step  $t$  we define an adaptive adversary that chooses  $G_{t+1}$  after observing  $V_t^{(0)}$  and  $V_t^{(1)}$ . The adversary chooses  $G_{t+1}$  such that the cut between  $V_t^{(0)}$  and  $V_t^{(1)}$  is of order of  $\Theta(\phi_t \cdot dn)$ . We show that such a graph exists when the number of nodes in both  $V_t^{(0)}$  and  $V_t^{(1)}$  is at least of linear size (in  $n$ ). By this choice the adversary ensures that the expected potential drop of  $\Psi(S_{t+1})$  at most  $-c\phi_t d/\Psi(s_t)$  for some constant  $c$ . Then we use the expected potential drop, together with the optional stopping theorem, to derive our lower bound.

We proceed by giving a lower bound on the potential drop assuming a cut-size of  $\Theta(\phi_t \cdot d \cdot n)$ .

**Lemma 2.5.** *Assume  $|\text{cut}(s_t, V \setminus s_t)| \leq c\phi_t \cdot dn$  for some constant  $c$ . Then we have*

$$\mathbf{E}[\Psi(S_{t+1}) \mid S_t = s_t] \geq \Psi(s_t) - \frac{c \cdot \phi_t \cdot d}{\Psi(s_t)}. \quad (7)$$

*Proof.* The proof is similar to the proof of Lemma 2.1.

$$\begin{aligned} \mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] &= \\ &= \mathbf{E}[\sqrt{\text{vol}(S_{t+1})} - \sqrt{\text{vol}(s_t)}] \\ &= \mathbf{E}\left[\sqrt{\min\{\text{vol}(V_{t+1}^{(0)}), m - \text{vol}(V_{t+1}^{(0)})\}} - \sqrt{\text{vol}(s_t)}\right] \end{aligned}$$

$$\begin{aligned}
&= \mathbf{E} \left[ \min \left\{ \sqrt{\text{vol}(v_t^{(0)}) + \Delta}, \sqrt{m - (\text{vol}(v_t^{(0)}) + \Delta)} \right\} - \sqrt{\text{vol}(v_t^{(0)})} \right] \\
&\geq \mathbf{E} \left[ \min \left\{ \sqrt{\text{vol}(v_t^{(0)}) + \Delta}, \sqrt{\text{vol}(v_t^{(0)}) - \Delta} \right\} - \sqrt{\text{vol}(v_t^{(0)})} \right] \\
&= \mathbf{E} \left[ \sqrt{\text{vol}(v_t^{(0)}) - |\Delta|} - \sqrt{\text{vol}(v_t^{(0)})} \right] \\
&= \mathbf{E} \left[ \sqrt{\text{vol}(v_t^{(0)})} \left( \sqrt{1 - \frac{|\Delta|}{\text{vol}(v_t^{(0)})}} - 1 \right) \right] \\
&\geq \Psi(s_t) \cdot \mathbf{E} \left[ \frac{|\Delta|}{2\Psi(s_t)^2} - \frac{|\Delta|^2}{\Psi(s_t)^4} \right]
\end{aligned} \tag{8}$$

where the last inequality comes from the Taylor expansion inequality  $\sqrt{1+x} \geq 1 + \frac{x}{2} - x^2$ ,  $x \geq -1$ . Similar to (2) we get

$$\begin{aligned}
\mathbf{E}[(\Delta)^2] &= \sum_{u \in V} \mathbf{Var}[X_u] + (\mathbf{E}[X_u])^2 = \sum_{u \in V} \mathbf{Var}[X_u] + 0 = \sum_{u \in V} (\mathbf{E}[X_u^2] - (\mathbf{E}[X_u])^2) \\
&\leq \sum_{u \in V} \mathbf{E}[X_u^2] = \sum_{u \in V} \frac{\lambda_u \cdot d}{2} = d \cdot |\text{cut}(s_t, V \setminus s_t)| \leq c \cdot \phi_t \cdot d \cdot \text{vol}(s_t) = c \cdot \phi_t \cdot d \Psi(s_t)^2.
\end{aligned}$$

From (8) we derive now

$$\mathbf{E}[\Psi(S_{t+1}) - \Psi(s_t) \mid S_t = s_t] \geq \Psi(s_t) \cdot \left( -\frac{c \cdot \phi_t \cdot d}{\Psi(s_t)^2} \right) \geq -\frac{c \cdot \phi_t \cdot d}{\Psi(s_t)}.$$

□

We now describe the adversary for the graphs that we use in our lower bound. We assume that we have initially two opinions and each of the two opinion is on  $n/2$  nodes. Recall that we can assume that in round  $t+1$  the adversary knows the graph  $G_t$  as well as the distribution of the opinions over the nodes. The adversary generates  $G_{t+1}$  as follows.  $c$  is a constant which is defined in Lemma A.3).

- If  $|s_t| \geq \gamma \cdot n$ , the adversary creates a  $d$ -regular graph  $G_{t+1}$  with two subsets  $s_t$  and  $V \setminus s_t$  such that the conductance of the  $\text{cut}(s_t, V \setminus s_t)$  is at most  $c \cdot \phi_t$ ; According to Lemma A.3 such a graph always exist.
- If  $|S_t| = |s_t| < \gamma \cdot n$ , the adversary does not change the graph and sets  $G_{t+1} = G_t$ .

To show Theorem 1.2 we first define a new potential function  $g$  and bound the one step potential drop. For  $x \geq 0$  we define  $g(x) = \frac{x^2}{2cd}$ . Since  $g(\cdot)$  is convex we obtain from Lemma 2.5, together with by Jensen's inequality that

$$\begin{aligned}
\mathbf{E}[g(\Psi(S_{t+1})) - g(\Psi(s_t)) \mid S_t = s_t] &\geq g(\mathbf{E}[\Psi(S_{t+1}) \mid S_t = s_t]) - g(\Psi(s_t)) \\
&\geq \frac{1}{2cd} \cdot \left( \left( \Psi(s_t) - \frac{c \cdot \phi_t \cdot d}{\Psi(s_t)} \right)^2 - (\Psi(s_t))^2 \right) \\
&\geq -\phi_t.
\end{aligned}$$

In the following lemma we use some Martingale arguments to show that with a probability of at least one half  $\text{vol}(S_{\tau''}) \geq \gamma d \cdot n/2$ . This implies that no opinion vanished after  $\tau''$  w.p.  $1/2$ , which yields Theorem 1.2.

**Lemma 2.6.** *Let  $|s_0| = n/2$ . Fix some constant  $\gamma < 1/4$ . Assume  $G_1, G_2, \dots$  is a sequence of graphs generated by the adversary defined above. Then*

$$P \left( g(\Psi(S_{\tau''})) > \frac{2\gamma \cdot n}{4c} \right) \geq \frac{1}{2}.$$

*Proof.* We define  $T' = \min\{\tau'', \min\{t : \text{vol}(S_t) \leq 2\gamma dn/2\}\}$ , where we recall that  $\tau''$  is the first rounds so that  $\sum_{t=1}^{\tau''} \phi_t \geq bn$ . We define

$$Z_t = g(\Psi(S_t)) + \sum_{i \leq t} \phi_i$$

and show it the following that  $Z_{t \wedge T'}$  is a sub-martingale with respect to the sequence  $S_1, S_2, \dots$ , where  $t \wedge T' = \min\{t, T'\}$ .

**Case  $t < T'$ :** For any  $t < T'$  we have

$$\begin{aligned}
E[Z_{t+1 \wedge T'} \mid S_t, \dots, S_1] &= E[Z_{t+1} \mid S_t, \dots, S_1] \\
&= \mathbf{E} \left[ g(\Psi(S_{t+1})) + \sum_{i \leq t+1} \phi_i \mid S_t, \dots, S_1 \right] \\
&= \mathbf{E} [g(\Psi(S_{t+1})) - g(\Psi(S_t)) \mid S_t, \dots, S_1] + g(\Psi(S_t)) + \sum_{i \leq t+1} \phi_i \\
&\geq -\phi_{t+1} + \left( g(\Psi(S_t)) + \sum_{i \leq t} \phi_i \right) + \phi_{t+1} \\
&= Z_t = Z_{t \wedge T'},
\end{aligned} \tag{9}$$

where the last equality follows since  $t < T'$ .

**Case  $t \geq T'$ :** For  $t \geq T'$  and we have

$$E[Z_{(t+1) \wedge T'} \mid S_t, \dots, S_1] = E[Z_{T'} \mid S_t, \dots, S_1] = Z_{T'} = Z_{t \wedge T'},$$

where the last equality follows since  $t \geq T'$ . Both cases together show that  $Z_{t \wedge T'}$  is a sub-martingale. According to Theorem 1.2 we have  $T' < \infty$ . Hence we can apply the optional stopping-time Theorem (c.f. Theorem A.4), which results in

$$E[Z_{t \wedge T'}] \geq E[Z_0] = \frac{n}{4c}.$$

We define

$$p = P \left( g(\Psi(S_{T'})) \leq \frac{2\gamma \cdot n}{4c} \right).$$

By definition of  $S_t$ , we have  $|S_t| \leq s_0$  and thus

$$g(\Psi(S_{T'})) \leq g(\Psi(s_0)) = \frac{n}{4c}.$$

Thus  $E[Z_{\tau''}] \leq \frac{n}{4c} + \sum_{i \leq \tau''} \phi_i$ . Hence we derive using  $T' \leq \tau''$  that

$$\begin{aligned}
\frac{n}{4c} &= E[Z_0] \leq E[Z_{T'}] \\
&\leq p \cdot \left( \frac{2\gamma \cdot n}{4c} + \sum_{i \leq T'} \phi_i \right) + (1-p) \cdot \left( \frac{n}{4c} + \sum_{i \leq \tau''} \phi_i \right) \\
&\leq 2p \cdot \gamma \cdot \frac{n}{4c} + (1-p) \cdot \frac{n}{4c} + \sum_{i \leq \tau''} \phi_i \\
&\leq 2p \cdot \gamma \cdot \frac{n}{4c} + (1-p) \cdot \frac{n}{4c} + b \cdot n + 1,
\end{aligned}$$

where the last inequality follows from the definition of  $\tau''$  together with the fact that  $\phi_i \leq 1$  for all  $i$ . Hence

$$0 \leq 2p \cdot \gamma \cdot \frac{n}{4c} - p \cdot \frac{n}{4c} + b \cdot n + 1$$

which equals

$$p \cdot (1 - 2\gamma) \cdot \frac{n}{4c} \leq b \cdot n + 1.$$

Thus for  $\gamma < 1/4$  and  $b < 1/(18c)$  we get  $p \leq 1/2$  and thus we have  $P(g(\Psi(S_{T'})) \leq \frac{2\gamma \cdot n}{4c}) = p \leq 1/2$  which yields the claim.  $\square$

In the following we observe that if the adversary is allowed to change the degrees, then the expected consensus time is super-exponential.

**Observation 1** (super exponential runtime). *Let  $G_1 = (V, E_1)$ ,  $G_2 = (V, E_2)$ ,  $\dots$  be a sequence of graphs with  $n$  nodes, where the edges  $E_1, E_2, \dots$  are distributed by an adaptive adversary, then the expected consensus time is at least  $\Omega((n/c)^{n/c})$  for some constant  $c$ .*

*Proof.* The idea is the following. The initial network is a line and there are two opinions 0 and 1 distributed on first  $n/2$  and last  $n/2$  nodes respectively. Whenever one opinion  $i$  has fewer nodes than the other, the adversary creates a graph where only one edge is crossing the cut between both opinions and the smaller opinions forms a clique. Hence, the probability for the smaller opinion to decrease is  $O(1/n)$  and the probability for the bigger opinion to decrease is  $\Omega(1)$ . This can be coupled with a biased random walk on a line of  $n/4$  nodes with the a transition probability at most  $1/(n/4) = 4/n$  in one direction and at least  $1/2$  in the other. Consequently, there exists a constant  $c$  such that expected consensus time is  $\Omega((n/c)^{n/c})$ .  $\square$

The following observation shows that the bound of Theorem 1.1 for static regular graphs of  $O(n/\phi)$  is tight for regular graphs if either the degree or the conductance is constant.

**Observation 2** (lower bound static graph). *For every  $n, d \geq 3$ , and constant  $\phi$ , there exists a  $d$ -regular graph  $G$  with  $n$  nodes and a constant conductance such that the expected consensus time on  $G$  is  $\Omega(n)$ . Furthermore, for every even  $n$ ,  $\phi > 1/n$ , and constant  $d$ , there exists a (static)  $d$ -regular graph  $G$  with  $\Theta(n)$  nodes and a conductance of  $\Theta(\phi)$  such that the expected consensus time on  $G$  is  $\Omega(n/\phi)$ .*

*Proof.* For the first part of the claim we argue that the meeting time of a  $d$ -regular graphs is  $\Omega(n)$  [2]. The claim follows from the duality between coalescing random walks and the voter model. For the second part of the claim we construct a random graph  $G'$  with  $n' = \Theta(n \cdot \phi)$  nodes and a constant conductance (See, e.g., [3]). We obtain  $G$  by replacing every edge  $(u, v)$  of  $G'$  with a path connecting  $u$  and  $v$  of length  $\ell = \Theta(1/\phi)$  and  $\ell \bmod d = 0$ . Additionally, we make  $G$   $d$ -regular by adding  $\ell(d-2)/d$  nodes to every path in such a way that the distance between  $u$  and  $v$  is maximised. We note that the obtained graph  $G$  has conductance  $\Theta(\phi)$ ,  $\Theta(n)$  nodes, and is  $d$ -regular. The meeting time of  $G'$  is  $\Omega(n')$  [2]. Taking into account that traversing every path in  $G$  takes in expectation  $\ell^2 = \Theta(1/\phi^2)$  rounds, the expected meeting time in  $G$  is at least  $\Omega(n' \cdot \ell^2) = \Omega(n/\phi)$ . This yields the claim.  $\square$

### 3 Analysis of the Biased Voter Model

In this section, we prove Theorem 1.3. We show that the set  $S_t$  of nodes with the preferred opinion grows roughly at a rate of  $1 + \Theta(\phi_t)$ , as long as  $S_t$  or  $S'_t$  has at least logarithmic size. For the analysis we break each round down into several steps, where exactly one node which has at least one neighbour in the opposite set is considered. Instead of analysing the growth of  $S_t$  for every round we consider larger time *intervals* consisting of a suitably chosen number of steps. We change the process slightly by assuming that there is always one node with the preferred opinion. If in some round the preferred opinions vanishes totally, Node 1 is set back to the preferred opinion. Symmetrically, if all other opinions vanishes, then Node 1 is set to Opinion 1. Note that this will only increase the runtime of the process.

The proof unfolds in the following way. First, we define formally the *step sequence*  $\mathcal{S}$ . Second, we define (Definition 1) a step sequence  $\mathcal{S}$  to be *good* if, intuitively speaking, the preferred opinion grows quickly enough in any sufficiently large subsequence of  $\mathcal{S}$ . Afterward, we show that if  $\mathcal{S}$  is a good step sequence, then the preferred opinion prevails in at most  $\tau'''$  rounds (Lemma 3.2). Finally, we show that that  $\mathcal{S}$  is indeed a good step sequence w.h.p. (Lemma 3.3).

We now give some definitions. Again, we denote by  $S_t$  the random set of nodes that have the preferred opinion right after the first  $t$  rounds, and let  $S'_t = V \setminus S_t$ . For a fixed time step  $t$  we write  $s_t$  and  $s'_t$ . We define the *boundary*  $\partial s_t$  as the subset of nodes in  $s'_t$  which are adjacent to at least one node from  $s_t$ . We use the symmetric definition for  $\partial s'_t$ . For each  $u \in V$ , let  $\lambda_{u,t}$  be the number of edges incident with  $u$  crossing the cut  $\text{cut}(s_t, s'_t)$ , or equivalently, the number of  $u$ 's neighbours that have a different opinion than  $u$ 's before round  $t$ .

We divide each round  $t$  into  $|s_t| + |s'_t|$  steps, in every step a single node  $v$  from either  $\partial s_t$  or  $\partial s'_t$  randomly chooses a neighbour  $u$  and adopts its opinion with the corresponding bias. Note that we assume that  $v$  sees  $u$ 's opinion referring to *beginning* of the round, even if was considered before  $v$  and changed its opinion in the meantime. It is convenient to label the steps independently of the round in which they take place. Hence, step  $i$  denotes the  $i$ -th step counted from the *beginning* of the first round. Also  $u_i$  refers to the node which considered in step  $i$  and  $\lambda_i = \lambda_{u_i,t}$ . We define the indicator variable  $o_i$  with  $o_i = 1$  if  $u_i$  has the preferred opinion and  $o_i = 0$  otherwise. Let

$$\Lambda(i) = \sum_{j=1}^i (1 - o_j) \cdot \lambda_j \quad \text{and} \quad \Lambda'(i) = \sum_{j=1}^i o_j \cdot \lambda_j.$$

Unfortunately, the order in which the nodes are considered in a round is important for our analysis and cannot be arbitrarily. Intuitively, we order the nodes in  $s_t$  and  $s'_t$  such that sum of the degrees of nodes which are already considered from  $s_t$  and the sum of the degrees of nodes already considered from  $s'_t$  differs by at most  $d$ , i.e.,

$$|\Lambda_i - \Lambda'_i| \leq d. \quad (10)$$

The following rule determines the node to be considered in step  $j + 1$ : if  $\Lambda(j) \leq \Lambda'(j)$ , then the (not yet considered) node  $v \in \partial s_t$  is with smallest identifier is considered. Otherwise the node  $v \in \partial s'_t$  is with smallest identifier is considered. Note that at the first step  $i$  of any round we have  $\Lambda_i = \Lambda'_i$ . This guarantees that (10) holds. The *step sequence*  $\mathcal{S}$  is now defined as a sequence of tuples, i.e.,  $\mathcal{S} = (u_1, Z_1), (u_2, Z_2), \dots$ , where  $Z_j = 1$  if  $u_j$  changed its opinion in step  $j$  and  $Z_j = 0$  otherwise for all  $j \geq 1$ . Observe that when given the initial assignment and the sequence up to step  $i$ , then we know the *configuration*  $\mathcal{C}_i$  of the system, i.e., the opinions of all nodes at step  $i$  and in which round step  $i$  occurred.

In our analysis we consider the increase in the number of nodes with the preferred opinion in time intervals which contain a sufficiently large number of steps, instead of considering one round after the other. The following definitions define these intervals.

For all  $i, k \geq 0$  where  $\mathcal{C}_i$  is fixed, we define the random variable  $S_{i,k} := \min\{j : \Lambda_j - \Lambda_i \geq k\}$ , which is the first time step such that nodes with a degree-sum of at least  $k$  were considered. Let  $I_{i,k} = [i + 1, S_{i,k}]$  be the corresponding interval where we note that the length is a random variable. We proceed by showing an easy observation.

**Observation 3.** *The number of steps in the interval  $I_{i,k}$  is at most  $2k + 2d$ , i.e.,  $|I_{i,k}| \leq 2k + 2d$ . Furthermore,  $\Lambda'(S_{i,k}) - \Lambda'(i) \leq k + 2d$ .*

*Proof.* Assume, for the sake of contradiction, that  $|I_{i,k}| > 2k + 2d$ . This implies, by definition of  $I_{i,k} = [i + 1, S_{i,k}]$ , that the number of edges of the preferred opinion considered in  $I = [i + 1, 2k + 2d]$  is strictly less than  $k$ , i.e.,  $\Lambda(i + 2k + 2d) - \Lambda(i) < k$ . We have

$$\begin{aligned} \Lambda'(i + 2k + 2d) - \Lambda(i + 2k + 2d) &> \Lambda'(i + 2k + 2d) - (\Lambda(i) + k) \\ &\geq \Lambda'(i) + 2k + 2d - k - (\Lambda(i) + k) \\ &= \Lambda'(i) - \Lambda(i) + 2d \\ &\geq -d + 2d \geq d, \end{aligned} \quad (11)$$

which contradicts (10). Hence  $|I_{i,k}| \leq 2k + 2d$ .

We now prove the second part of the lemma. Assume, for the sake of contradiction,  $\Lambda'(S_{i,k}) - \Lambda'(i) > k + 2d$ . This implies that

$$\begin{aligned} \Lambda'(S_{i,k}) - \Lambda(S_{i,k}) &> \Lambda'(i) + k + 2d - \Lambda(S_{i,k}) \\ &\geq \Lambda'(i) + k + 2d - (\Lambda(i) + k) \\ &= \Lambda'(i) - \Lambda(i) + 2d \\ &\geq -d + 2d \geq d, \end{aligned} \quad (12)$$

which contradicts (10). Hence  $\Lambda'(S_{i,k}) - \Lambda'(i) \leq k + 2d$ .  $\square$

Fix  $\mathcal{C}_i$  and let  $X_{i,k}$  be the total number of times during interval  $I_{i,k}$  that a switch from a non-preferred opinion to the preferred one occurs; and define  $X'_{i,k}$  similarly for the reverse switches. Finally, we define  $Y_{i,k} = X_{i,k} - X'_{i,k}$ ; thus  $Y_{i,k}$  is the increase in number of nodes that have the preferred opinion during the time interval  $I_{i,k}$ .

Define  $\ell = \frac{132\beta \log n}{(1-\alpha_1)^2}$  and  $\beta' = \frac{600d}{\alpha_1 \cdot (1-\alpha_1)^2}$ . In the following we define a *good* sequence.

**Definition 1.** *We call the sequence  $\mathcal{S}$  of steps good if it has all of the following properties for all  $i \leq T' = 2\beta' \cdot n$ . Consider the first  $T'$  steps of  $\mathcal{S}$  (fix  $\mathcal{C}_{T'}$ ). Then,*

- (a)  $Y_{0,T'} \geq 2n$ . (The preferred opinion prevails in at most  $T'$  steps)
- (b)  $Y_{0,i} + |S_0| > 1$ . (The preferred opinion never vanishes)
- (c) For any  $1 \leq k \leq T'$  we have  $Y_{i,k} \geq -\ell$ . (# nodes of the pref. opinion never drops by  $\ell$ )
- (d) For any  $\ell \leq \gamma \leq T'$ , we have  $Y_{i,k} > \gamma$ , where  $k = \gamma \cdot \beta'$ . (# nodes of the pref. opinion increases)

This definition allows us to prove in a convenient way that a step sequence  $\mathcal{S}$  is w.h.p. good: For each property, we simply consider each (sufficiently large) subsequence  $S$  separately and we show that w.h.p.  $S$  has the desired property. We achieve this by using a concentration bound on  $Y_{i,k}$  which we establish in Lemma 3.1. Afterward, we take union bound over all of these subsequences and properties. Using the union bound allows us to show the desired properties in all subsequences in spite of the emerging dependencies. This is done in Lemma 3.2.

We now show the concentration bounds on  $Y_{i,k}$ . These bounds rely on the Chernoff-type bound established in Lemma A.5. This Chernoff-type bound shows concentration for variables having the property that the sum of the conditional probabilities of the variables, given all previous variables, is always bounded (from above or below) by some  $b$ . The bound might be of general interest.

**Lemma 3.1.** *Fix configuration  $\mathcal{C}_i$ . Then,*

(a) *For  $k = \gamma \frac{256d}{\alpha_1 \cdot (1-\alpha_1)^2}$  with  $\gamma \geq 1$  it holds that*

$$\Pr(Y_{i,k} < \gamma) \leq \exp(-\gamma).$$

(b) *For  $k \geq 0$ , any  $b' = \alpha_1 \cdot (k + 2d)/d$ , and any  $\delta > 0$  it holds that*

$$\Pr(Y_{i,k} < -(1 + \delta)b') \leq \exp\left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{b'}.$$

*Proof.* First we show a lower bound for  $X_{i,k}$  and an upper bound for  $X'_{i,k}$ .

For  $1 \leq j \leq 2k + 2d$ , let  $A_j$  be the event that  $u_{i+j}$  switches its opinion from any opinion other than the preferred opinion to the preferred one. Furthermore, for  $1 \leq j \leq 2k + 2d$ , let

$$Z_j = \begin{cases} 1 & \text{if } i+j \leq S_{i,k} \text{ and } A_j \\ 0 & \text{otherwise} \end{cases}$$

Define

$$P_j = \begin{cases} \Pr(Z_j = 1 \mid Z_1, \dots, Z_{j-1}) = \lambda_{i+j}/d & i+j \leq S_{i,k} \text{ and } A_j \\ 0 & \text{otherwise} \end{cases}$$

It follows that  $X_{i,k} = \sum_j Z_j$ . We set  $B = \sum_j P_i$ . We derive

$$B = \sum_j P_i = \sum_{s \in I_{i,k}} (1 - o_s) \lambda_s / d = (\Lambda(S_{i,k}) - \Lambda(i)) / d \geq k/d,$$

by definition of  $I_{i,k}$ . From Lemma A.5(b) with  $b^* = k/d \leq B$  we obtain for any  $0 < \delta < 1$ , that

$$\Pr(X_{i,k} < (1 - \delta)b^*) < e^{-b^* \cdot \delta^2 / 2} \quad (13)$$

We now use a similar reasoning to bound  $X'_{i,k}$ . For  $1 \leq j \leq 2k + 2d$  let  $A'_j$  be the event that  $u_{i+j}$  switches its opinion from the preferred opinion to any other opinion. Define for  $1 \leq j \leq 2k + 2d$  the variables  $Z'_j$  similar to  $Z_j$  but using  $A'_j$  instead of  $A_j$ . And similarly, define for  $1 \leq j \leq 2k + 2d$  the variables  $P'_j$  similar to  $P_j$  but using  $Z'_j, \dots, Z'_1$  instead of  $Z_j, \dots, Z_1$ . In the same spirit as before one can define  $B' = \sum_j P'_i$ . We derive

$$B' = \sum_j P'_i \leq \sum_{s \in I_{i,k}} o_s \cdot \alpha_1 \cdot \lambda_s / d = \alpha_1 \cdot (\Lambda'(S_{i,k}) - \Lambda'(i)) / d \leq \alpha_1 \cdot (k + 2d) / d,$$

where the last inequality follows from Observation 3.

Using a similar reasoning for  $X'_{i,k}$  and applying Lemma A.5(a) with  $b' = \frac{\alpha_1 \cdot (k+2d)}{d} \geq B'$ , yields for  $\delta > 0$ ,

$$\Pr(X'_{i,k} > (1 + \delta)b') < \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{b'} \quad (14)$$

To prove part (a) we now combine (13) and (14) as follows. Since  $k \geq 16d/(1 - \alpha_1)$ , for  $\delta = (1 - \alpha_1)/8$ , we have that w.p. at least  $1 - e^{-b^* \cdot \delta^2 / 2} - \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{b'}$  that

$$\begin{aligned}
Y_{i,k} &= X_{i,k} - X'_{i,k} \\
&\geq (1-\delta)b^* - (1+\delta)b' \\
&\geq \frac{k}{d} \left( 1-\delta - (1+\delta)\alpha_1 - (1+\delta)\frac{2d}{k} \right) \\
&\geq \frac{k}{d} \left( 1-\alpha_1 - 2\delta - (1+\delta)\frac{2d}{k} \right) \\
&\geq \frac{k}{d} \left( 1-\alpha_1 - (1-\alpha_1)/4 - (1+\delta)\frac{2d}{k} \right) \\
&\geq \frac{k}{d} (1-\alpha_1 - (1-\alpha_1)/4 - (1-\alpha_1)/4) \\
&\geq (1-\alpha_1)\frac{k}{2d} \\
&= \gamma \frac{32}{\alpha_1 \cdot (1-\alpha_1)} \\
&> \gamma.
\end{aligned} \tag{15}$$

We proceed by bounding  $1 - e^{-b^* \cdot \delta^2/2} - \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{b'}$ . Let  $p = e^{-b^* \cdot \delta^2/2} + \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{b'}$ . We have

$$\begin{aligned}
p &\leq e^{-b^* \cdot \delta^2/2} + \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{b'} \\
&\leq e^{-b^* \cdot \delta^2/2} + e^{-b' \cdot \delta^2/2} \\
&= e^{-b^* \cdot (1-\alpha_1)^2/128} + e^{-b' \cdot (1-\alpha_1)^2/128} \\
&\leq 2e^{-\alpha_1 \cdot (1-\alpha_1)^2 \cdot k/(128d)} \\
&\leq e^{-\alpha_1 \cdot (1-\alpha_1)^2 \cdot k/(266d)} \\
&= e^{-\gamma},
\end{aligned} \tag{16}$$

where we used that  $k \geq \frac{256d}{\alpha_1 \cdot (1-\alpha_1)^2}$ . Part (a) follows from (15) and (16).

We now prove part (b). Since we used that  $Y_{i,k} \geq -X'_{i,k}$ , the bound from (14) implies

$$\Pr(Y_{i,k} < -(1+\delta)b') \leq \Pr(X'_{i,k} > (1+\delta)b') < \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{b'}$$

Hence, part (b) of the lemma follows. This completes the proof of Lemma 3.1. □

The following two lemmas imply the theorem.

**Lemma 3.2.** *Let  $\mathcal{S}$  be a step sequence. Then  $\mathcal{S}$  is good with high probability.*

*Proof.* We show for every property of Definition 1 that it holds with high probability. Fix an arbitrary  $i \leq T'$ .

(a) We derive from Lemma 3.1(a) with  $k = 2n \cdot \beta'$  that  $\Pr(Y_{0,k} < 2n) \leq e^{-2n} \leq n^{-\beta}$ .

(b) Follows from (c).

(c) Fix an arbitrary  $k \leq T'$ . We distinguish between two cases

(i) Case  $k \leq \beta \cdot \log n \cdot \beta'$ . We derive from Lemma 3.1(b) the following. For  $k \geq 0$ , any  $b' = \alpha_1 \cdot (k + 2d)/d \leq \frac{300\beta \log n}{(1-\alpha_1)^2} = \ell/2$ , and any  $\delta > 0$  it holds that

$$\Pr(Y_{i,k} < -(1+\delta)b') \leq \exp \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{b'}.$$

We distinguish once more between two cases.

- If  $\sqrt{\frac{3\beta \log n}{b'}} < 1$  set  $\delta = \sqrt{\frac{3\beta \log n}{b'}} < 1$ . We have

$$\begin{aligned} \Pr(Y_{i,k} < -\ell) &\leq \Pr(Y_{i,k} < -2b') \leq \Pr(Y_{i,k} < -(1+\delta)b') \\ &\leq \exp\left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{b'} \leq \exp(-\delta^2 b'/3) \leq n^{-\beta}. \end{aligned}$$

- Otherwise, we have  $b' \leq 3\beta \log n < 4\beta \log n$ . Set  $\delta = \frac{4\beta \log n}{b'} > 1$ . We have

$$\Pr(Y_{i,k} < -\ell) \leq \Pr(Y_{i,k} < -(1+\delta)b') \leq \exp\left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{b'} \leq \exp(-\delta b'/3) \leq n^{-\beta}.$$

Thus, the claim follows.

- (ii) Case  $k > \beta \cdot \log n \cdot \beta'$ . We derive from Lemma 3.1(a) that  $\Pr(Y_{i,k} < -\ell) \leq \Pr(Y_{i,k} < \beta \log n) < n^{-\beta}$ .

Hence, in all cases we have  $\Pr(Y_{i,k} < -\ell) \leq n^{-\beta}$ .

- (d) We derive from Lemma 3.1(a) with  $k = z \cdot \beta'$  that  $\Pr(Y_{i,k} < z) < \exp(-z) \leq n^{-\beta}$ .

Since  $0 < \alpha_1 < 1$ , we have that  $T' \leq n^3$ . The number of events in Definition 1 is bounded by  $5T'^2 \leq n^7$ . Thus choosing,  $\beta \geq 8$ , and taking union bound over all these events yields the claim.  $\square$

**Lemma 3.3.** *If  $\mathcal{S}$  is a good step sequence, then in at most  $T'$  time steps, the preferred opinion prevails and the  $T'$  time steps occur before round  $\tau'''$ .*

*Proof.* Recall, that we assume that if in some round the preferred opinions vanishes, then after this round, the opinion of some fixed node switches spontaneously. Similarly, if in some round the preferred opinion prevails, then after this round, the opinion of some fixed node switches spontaneously to an arbitrary other opinion. the preferred opinion never vanishes. This process  $P'$  diverges from the original process  $P$  only after the first step where either the preferred opinion prevails or vanishes. From (a) and (b) of the definition of a good sequence it follows that the preferred opinion prevails in  $P'$  after  $T'$  steps. It is easy to couple both process so that the good opinion also prevails in the original process  $P$ .

It remains to argue that the  $T'$  time steps happen before round  $\tau'''$ . Using the definition of the conductance, we can lower bound the number of steps in any round  $t$  by  $|cut(S_t, S'_t)| \geq d \cdot \min\{|S_t|, |S'_t|\} \cdot \phi_t$ . We then consider intervals of sufficient length in which the size of the preferred doubles as long as its size is below  $n/2$ . Afterward, one can argue that size of all the non preferred opinions halves every interval. We now give some intuition for the remainder of the proof. Consider the following toy case example of a static graphs with  $\alpha_1 = 0$  (rumor spreading). The length of an interval required for the preferred  $S$  with  $|s| \leq n/3$  to double in expectation is bounded by  $1/\phi$ . In our setting, we need to handle two main difference w.r.t. the toy case example. First, the number of nodes with the preferred opinion can reduce by up to  $\beta \log n$  (Definition 1(c)). Since  $\beta$  is constant, this can be easily compensated by slightly longer intervals. Second, the graph is dynamic as opposed to static. To address this we 'discretize', similarly as before, rounds into consecutive phases which ensure that sum of the  $\phi_t$  for rounds  $t$  in this phase is at least 1. Thus, in our toy example one requires 1 phases in expectation.

We proceed by discretizing the rounds into phases. Phase  $i$  starts at round  $\tau(i) = \min\{t : \sum_{j=1}^t \phi_j \geq 2i\}$  for  $i \geq 0$  and it ends at round  $\tau(i+1) - 1$ . Since  $\phi_j \leq 1$  for all  $j \geq 0$  we have  $\tau(0) < \tau(1) < \dots$  and  $\sum_{j=\tau(i)}^{\tau(i+1)-1} \phi(j) \geq 1$  for  $i \geq 0$ . We now map the steps of  $\mathcal{S}$  to rounds. For this we define the *check point*  $t_j$  to be the following round for  $j \leq j_{max} = 4 \log n + 1$ .

$$t_j = \begin{cases} 0 & \text{if } j = 0 \\ \tau(12\ell \cdot \beta'/d) & \text{if } j = 1 \\ \tau(12\ell \cdot \beta'/d + (j-1) \cdot 24\beta'/d) & \text{if } 2 \leq j \leq 4 \log n \\ \tau(24\ell \cdot \beta'/d + (j_{max}-1) \cdot 24\beta'/d) & \text{if } j = j_{max} \end{cases}$$

Given any good sequence  $\mathcal{S}$ , we show by induction over  $j$  that the following lower bounds on the size of the preferred opinion at these check points. More concretely, define for all  $j \leq j_{max}$  that



$$\zeta(j) = \begin{cases} 0 & \text{if } j = 0 \\ 2\ell & \text{if } j = 1 \\ \min\{2\ell \cdot 2^{j-2}, n/2\} & \text{if } 2 \leq j \leq 2 \log n \\ \min\{n - 2^{\log(n/2) - (j-2 \log n)}, n - 2\ell\} & \text{if } 2 \log n < j \leq 4 \log n \\ n & \text{if } j = j_{\max} \end{cases}$$

We now show for all  $j \leq j_{\max}$  that

$$|s_{t_j}| \geq \zeta(j), \quad (17)$$

We consider each of the cases depending of the size of  $j$  w.r.t. (17). The induction hypothesis  $j = 0$  holds trivially. Suppose the claim holds for  $j - 1$  for  $j \geq 1$ .

We assume w.l.o.g. the following about  $\mathcal{S}$ : there is no step  $t$  before round  $t_j$  with  $|s_t| \geq \zeta(j) + \ell$  since by Definition 1(c) this implies that  $|s_{t_j}| \geq \zeta(j)$  which yields the inductive step. This assumption, implies that for all  $t \leq t_j$  we have  $|s_t| < |\zeta(j)| + \ell$ . On the other hand, by induction hypothesis and Definition 1(c) we have  $\zeta(j-1) - \ell \leq |s_t|$ . Thus, we assume in the following

$$|s_t| \in [\zeta(j-1) - \ell, \zeta(j) + \ell] \quad (18)$$

We now distinguish between the following cases based on  $j$ .

- $j = 1$  : In every step  $t \in (0, t_1]$  the number of edges crossing the cut is at least  $\phi_t \cdot d$  and hence the number of edges crossing the cut in the interval  $(0, t_1]$  is at least  $\sum_{i=1}^{t_1} \phi_i \cdot d \geq 12\ell \cdot \beta'$ . Let  $k = 3\ell \cdot \beta'$ . Definition 1(d) implies that  $Y_{1,k} > 3\ell$ . Hence, by Definition 1(c) we have  $|s_{t_1}| \geq \zeta(1)$  as desired.
- $2 \leq j \leq 2 \log n$  : In every step  $t \in (t_{j-1}, t_j]$  we have, by (18), that the number of edges crossing the cut is at least

$$\phi \cdot d \cdot (\min\{|s_t|, |s'_t|\} - \ell) \geq \phi \cdot d \cdot \min\{\zeta(j-1) - \ell, n/2 - \ell\} \geq \phi \cdot d \cdot (\zeta(j-1) - \ell) \geq \phi_t \cdot d \cdot \zeta(j-1)/2.$$

Hence the number of edges crossing the cut in the interval  $(t_{j-1}, t_j]$  is at least  $12\zeta(j-1) \cdot \beta'$ . Let  $k = 3\zeta(j-1) \cdot \beta'$ . Definition 1(d) implies that  $Y_{t_{j-1}, k} \geq 3\zeta(j-1) \geq \zeta(j) + \ell$ . Hence, by Definition 1(c) we have  $|s_{t_j}| \geq \zeta(j)$ .

- $2 \log n < j \leq 2 \log n$  : In every step  $t \in (t_{j-1}, t_j]$  we have, by (18), that the number of edges crossing the cut is at least

$$\phi_t \cdot d \cdot (\min\{|s_t|, |s'_t|\} - \ell) \geq \phi \cdot d \cdot (n - \zeta(j) - \ell) \geq \phi \cdot d \cdot (n - \zeta(j))/2.$$

Hence the number of edges crossing the cut in the interval  $(t_{j-1}, t_j]$  is at least  $\sum_{i=t_{j-1}+1}^{t_j} \phi_i \cdot d \cdot (n - \zeta(j))/2 \geq 12(n - \zeta(j)) \cdot \beta'$ . Let  $k = 3(n - \zeta(j)) \cdot \beta'$ . Definition 1(d) implies that  $Y_{t_{j-1}, k} \geq 3(n - \zeta(j)) \geq (n - \zeta(j)) + \ell$ . Hence, by Definition 1(c) we have  $|s_{t_j}| \geq \min\{\zeta(j-1) + (n - \zeta(j)), n - 2\ell\} \geq \zeta(j)$ .

- $j_{\max}$  : In every step  $t \in (t_{j-1}, t_j]$  we have, by (18), that the number of edges crossing the cut at any time step is at least  $\phi_t \cdot d$ . Hence the number of edges crossing the cut in the interval  $(t_{j-1}, t_j]$  is at least  $d \geq 12\ell \cdot \beta'$ . Let  $k = 3\ell \cdot \beta'$ . Definition 1(d) implies that  $Y_{t_{j_{\max}-1}, k} \geq 3\ell$ . Hence, by Definition 1(c) we have  $|s_{t_{j_{\max}}}| \geq n = \zeta(j_{\max})$ .

This completes the proof of (17). We have

$$t_{j_{\max}} = \tau(24\ell \cdot \beta' + j_{\max} \cdot 24\beta'/d) \leq 4(24\ell \cdot \beta' + (j_{\max} - 1) \cdot 24\beta'/d) \leq \tau''',$$

which yields the proof. □

*Proof of Theorem 1.3.* The claim follows from Lemma 3.2 together with Lemma 3.3. □

## References

- [1] M. A. Abdullah and M. Draief. Global majority consensus by local majority polling on graphs of a given degree sequence. *Discrete Applied Mathematics*, 180:1–10, 2015.
- [2] D. Aldous and J. Fill. Reversible markov chains and random walks on graphs, 2002. Unpublished, <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [3] N. Alon, O. Schwartz, and A. Shapira. An elementary construction of constant-degree expanders. *Comb. Probab. Comput.*, 17(3):319–327, May 2008.
- [4] C. Avin, M. Koucký, and Z. Lotker. How to explore a fast-changing world (Cover time of a simple random walk on evolving graphs). In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 121–132, 2008.
- [5] Y. Azar, A. Broder, A. Karlin, and E. Upfal. Balanced allocations. *SIAM J. Comput.*, 29(1):180–200, 1999.
- [6] C. Cooper, R. Elsässer, H. Ono, and T. Radzik. Coalescing random walks and voting on connected graphs. *SIAM J. Discrete Math.*, 27(4):1748–1758, 2013.
- [7] C. Cooper, R. Elsässer, and T. Radzik. The power of two choices in distributed voting. In *Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP)*, pages 435–446, 2014.
- [8] C. Cooper, R. Elsässer, T. Radzik, N. Rivera, and T. Shiraga. Fast consensus for voting on general expander graphs. In *Proceedings of the 29th International Symposium on Distributed Computing (DISC)*, pages 248–262, 2015.
- [9] J. Cruise and A. Ganesh. Probabilistic consensus via polling and majority rules. *Queueing Systems*, 78(2):99–120, 2014.
- [10] J. Díaz, L. Goldberg, G. Mertzios, D. Richerby, M. Serna, and P. Spirakis. Approximating fixation probabilities in the generalized moran process. *Algorithmica*, 69(1):78–91, 2014.
- [11] P. Donnelly and D. Welsh. Finite particle systems and infection models. *Mathematical Proceedings of the Cambridge Philosophical Society*, 94:167–182, 1983.
- [12] R. Elsässer, T. Friedetzky, D. Kaaser, F. Mallmann-Trenn, and H. Trinker. Efficient  $k$ -party voting with two choices. *ArXiv e-prints*, Feb. 2016.
- [13] G. Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *Proceedings of the 28th International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 57–68, 2011.
- [14] G. Giakkoupis, T. Sauerwald, and A. Stauffer. Randomized rumor spreading in dynamic graphs. In *Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP)*, pages 495–507, 2014.
- [15] Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Information and Computation*, 171(2):248–268, 2001.
- [16] R. Holley and T. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability*, 3(4):643–663, 1975.
- [17] B. Houchmandzadeh and M. Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics*, 13(7):073020, 2011.
- [18] M. Kearns and J. Tan. Biased voting and the democratic primary problem. In *Proceedings of the 4th International Workshop on Internet and Network Economics (WINE)*, pages 639–652, 2008.
- [19] F. Kuhn, N. Lynch, and R. Oshman. Distributed computation in dynamic networks. In *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC)*, pages 513–522, 2010.

- [20] H. Lam, Z. Liu, M. Mitzenmacher, X. Sun, and Y. Wang. Information dissemination via random walks in  $d$ -dimensional space. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1612–1622, 2012.
- [21] N. Lanchier and C. Neuhauser. Voter model and biased voter model in heterogeneous environments. *Journal of Applied Probability*, 44(3):770–787, 2007.
- [22] T. Liggett. *Interacting Particle Systems*. Springer Berlin Heidelberg, 1985.
- [23] G. Mertzios and P. Spirakis. Strong bounds for evolution in networks. In *Proceedings of the 40th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 669–680, 2013.
- [24] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge Univ. Press, 2005.
- [25] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.
- [26] T. Nakata, H. Imahayashi, and M. Yamashita. A probabilistic local majority polling game on weighted directed graphs with an application to the distributed agreement problem. *Networks*, 35(4):266–273, 2000.
- [27] R. Oliveira. On the coalescence time of reversible random walks. *Transactions of the American Mathematical Society*, 364(4):2109–2128, 2012.
- [28] Y. Peres, A. Sinclair, P. Sousi, and A. Stauffer. Mobile geometric graphs: Detection, coverage and percolation. *Probability Theory and Related Fields*, 156(1-2):273–305, 2013.

## A Auxiliary Claims

**Lemma A.1.** Let  $X_i$  and  $Y_i$  be the random variables defined in the proof of Lemma 2.1. Let  $f(\cdot)$  be a concave and continuous function. We have

$$E[f(\sum_i X_i)] \leq E[f(\sum_i Y_i)].$$

*Proof.* We show by induction over  $i$  that the variables

$$E[f(Y_1 + \dots + Y_{i-1} + X_i + \dots + X_k)] \leq E[f(Y_1 + \dots + Y_i + X_{i+1} + \dots + X_k)].$$

Let  $Z = Y_1 + \dots + Y_{i-1} + X_{i+1} + \dots + X_k$ . In step  $i \rightarrow i+1$  we have

$$\begin{aligned} E[f(Y_1 + \dots + Y_{i-1} + X_i + \dots + X_k)] &= E[f(Z + X_i)] \\ &= \frac{\lambda_i}{d_i} E[f(Z + d_i)] + \left(1 - \frac{\lambda_i}{d_i}\right) E[f(Z)] \\ &= \lambda_i E\left[\frac{f(Z + d_i)}{d_i}\right] + \left(1 - \frac{\lambda_i}{d_i}\right) E[f(Z)] \\ &= \lambda_i E\left[\frac{f(Z + d_i) - f(Z)}{d_i}\right] + E[f(Z)] \\ &\leq \lambda_i E\left[\frac{f(Z + \lambda_i) - f(Z)}{\lambda_i}\right] + E[f(Z)] \\ &= E[f(Y_1 + \dots + Y_i + X_{i+1} + \dots + X_k)], \end{aligned}$$

where the last inequality follows from the concavity of  $f(\cdot)$ .  $\square$

**Lemma A.2.** Let  $Z_1, \dots, Z_n$  be independent random variables and  $Z = \sum_i Z_i$ . If  $\mathbf{E}[Z] = 0$ , then  $\mathbf{E}[Z^3] = \sum_i (\mathbf{E}[Z_i^3] - 3\mathbf{E}[Z_i^2] \cdot \mathbf{E}[Z_i] + 2\mathbf{E}[Z_i]^3)$ .

*Proof.* In the following we make use of  $\mathbf{E}[\sum_i Z_i] = 0$ . We derive

$$\begin{aligned} \mathbf{E}[Z^3] &= \mathbf{E}\left[\sum_{i,j,k} Z_i Z_j Z_k\right] \\ &= \sum_i \mathbf{E}\left[Z_i \sum_{j,k} Z_j Z_k\right] \\ &= \sum_i \mathbf{E}[Z_i^3] + 3 \sum_i \mathbf{E}[Z_i^2] \mathbf{E}\left[\sum_{j, j \neq i} Z_j\right] \\ &\quad + \sum_i \mathbf{E}[Z_i] \sum_{j, j \neq i} \mathbf{E}[Z_j] \mathbf{E}\left[\sum_{k, k \neq i, j} Z_k\right] \\ &= \sum_i \mathbf{E}[Z_i^3] + 3 \sum_i \mathbf{E}[Z_i^2] (0 - \mathbf{E}[Z_i]) \\ &\quad + \sum_i \mathbf{E}[Z_i] \sum_{j, j \neq i} \mathbf{E}[Z_j] (0 - \mathbf{E}[Z_i] - \mathbf{E}[Z_j]) \\ &= \sum_i \mathbf{E}[Z_i^3] - 3 \sum_i \mathbf{E}[Z_i^2] \mathbf{E}[Z_i] + \sum_i \mathbf{E}[Z_i] (-\mathbf{E}[Z_i]) \sum_{j, j \neq i} \mathbf{E}[Z_j] \\ &\quad + \sum_i \mathbf{E}[Z_i] \sum_{j, j \neq i} \mathbf{E}[Z_j] (-\mathbf{E}[Z_j]) \\ &= \sum_i \mathbf{E}[Z_i^3] - 3 \sum_i \mathbf{E}[Z_i^2] \mathbf{E}[Z_i] + 2 \sum_i \mathbf{E}[Z_i] (-\mathbf{E}[Z_i]) (-\mathbf{E}[Z_i]) \\ &= \sum_i \mathbf{E}[Z_i^3] - 3 \sum_i \mathbf{E}[Z_i^2] \mathbf{E}[Z_i] + 2 \sum_i \mathbf{E}[Z_i]^3. \end{aligned}$$

$\square$

We now show  $d$ -regular graphs with a cut of size  $\Theta(\phi_t dn)$  exist indeed.

**Lemma A.3.** *Let  $\frac{1}{nd} \leq \phi \leq 1$ . Let  $0 < \gamma < 1$  be some constant. Let  $d \geq 6$  be an even integer. For any integer  $n' \in [\gamma n, n/2]$  there exists a  $d$ -regular graph  $G = (V, E)$  with  $n$  nodes and the following property. There is a set  $S \subset V$  with  $|S| = n'$  such that  $|\text{cut}(S, V \setminus S)| = \Theta(\phi dn)$ . Moreover, there are at least  $n'/2$  nodes without any edges in  $\text{cut}(S, V \setminus S)$ .*

*Proof.* In the following we create two  $d$ -regular graphs  $G'$  and  $G''$  and connect them to a  $d$ -regular graph  $G$  such that the cut size is  $\Theta(\phi dn)$ . Let  $G' = (V', E')$  be the circulant graph  $C_{n'}^{\lfloor d/2 \rfloor}$  with  $V' = \{v'_1, \dots, v'_{n'}\}$ . Let  $G'' = (V'', E'')$  be the circulant graph  $C_{n-n'}^{\lfloor d/2 \rfloor}$  with  $V'' = \{v''_1, \dots, v''_{n-n'}\}$ . We now connect  $G'$  and  $G''$ . W.l.o.g.  $\phi \leq 1/d$ . The case  $\phi > 1/d$  is analogue. We choose  $k$  such that we have  $2 \leq k < n'/2$  and  $k = \Theta(\phi dn')$ . Let  $S' = \{v'_1, \dots, v'_k\}$  and let  $S'' = \{v''_1, \dots, v''_k\}$ . Now we remove all edges  $(v'_i, v'_{i+1})$  with  $1 \leq i \leq k-1$  and  $(v''_i, v''_{i+1})$ . Note that  $G'$  ( $G''$  respectively) is still connected. Furthermore, the vertices  $v'_1, v'_k, v''_1$  and  $v''_k$  have degree  $d-1$  and all other vertices of  $S' \cup S''$  have degree  $d-2$ . One can easily add  $(2k-2)$  edges such that (i) one endpoint of every edge is in  $V'$  and one in  $V''$ , and (ii) all vertices in  $G'$  and  $G''$  have degree  $d$ .

Note that the obtained graph  $G$  is connected and the cut  $|\text{cut}(V', V \setminus V')|$  contains  $(2k-2) = \Theta(\phi dn')$  edges. The claim follows directly.  $\square$

**Theorem A.4** ([24, Theorem 12.2]). *If  $Z_0, Z_1, \dots$  is a martingale with respect to  $X_1, X_2, \dots$  and if  $T$  is a stopping time for  $X_1, X_2, \dots$ , then*

$$E[Z_t] = E[Z_0]$$

whenever one of the following holds:

- The  $Z_i$  are bounded, so there is a constant for all  $i$ ,  $|Z_i| \leq c$ ;
- $T$  is bounded;
- $E[T] \leq \infty$ , and there is a constant  $c$  such that  $E[Z_{t+1} - Z_t | X_1, \dots, X_t] < c$ ;

## A.1 A Chernoff-type bound

The following lemma bounds the sum of (dependent) binary random variables, under the assumption that the sum of the conditional probabilities of the variables, given all previous variables, is always bounded (from above or below) by some  $b$ . The bounds are the same as the ones for independent variables but use  $b$  in place of  $\mu$ . The bound can be seen as a generalisation of [5]. The proof follows the proof of the independent case.

**Lemma A.5** (Chernoff Bound for Dependent Setting). *Let  $Z_1, Z_2, \dots, Z_\ell$  be a sequences of binary random variables, and for each  $1 \leq i \leq \ell$ , let  $p_i = \Pr(Z_i = 1 | Z_1, \dots, Z_{i-1})$ . Let  $Z = \sum_i Z_i$ ,  $B = \sum_i p_i$ . (a) For any  $b \geq 0$  with  $\Pr(B \leq b) = 1$ , it holds for any  $\delta > 0$  that*

$$\Pr(Z > (1 + \delta) \cdot b) < \left( \frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^b.$$

(b) For any  $b \geq 0$  with  $\Pr(B \geq b) = 1$ , then for any  $0 < \delta < 1$  it holds that

$$\Pr(Z < (1 - \delta)b) < e^{-b\delta^2/2}.$$

*Proof.* The proof follows the proof of the Chernoff bound given in [25]. However, the random variables  $\{Z_i : 1 \leq i \leq k\}$  are not independent. For any positive real  $t$  we have  $\Pr(Z \geq (1 + \delta)b) = \Pr(\exp(tZ) \geq \exp(t(1 + \delta)b))$ . Thus, by applying Markov inequality we derive

$$\Pr(Z \geq (1 + \delta)b) < \frac{\mathbf{E}[\exp(tZ)]}{\exp(t(1 + \delta)b)}. \quad (19)$$

We now bound  $\mathbf{E}[\exp(tZ)]$ . By law of total expectation and  $p_1 = \Pr(Z_1 = 1)$  we get

$$\begin{aligned} \mathbf{E}[\exp(tZ)] &= \mathbf{E}[\exp(tZ) | Z_1 = 1] \Pr(Z_1 = 1) + \mathbf{E}[\exp(tZ) | Z_1 = 0] \Pr(Z_1 = 0) \\ &= \mathbf{E}[\exp(tZ) | Z_1 = 1] p_1 + \mathbf{E}[\exp(tZ) | Z_1 = 0] (1 - p_1) \end{aligned}$$

$$\begin{aligned}
&= \mathbf{E} \left[ \exp \left( t \sum_{i=2}^k Z_i \right) \middle| Z_1 = 1 \right] \exp(t)p_1 + \mathbf{E} \left[ \exp \left( t \sum_{i=2}^k Z_i \right) \middle| Z_1 = 0 \right] (1 - p_1) \\
&\leq \max \left\{ \mathbf{E} \left[ \exp \left( t \sum_{i=2}^k Z_i \right) \middle| Z_1 = 1 \right], \mathbf{E} \left[ \exp \left( t \sum_{i=2}^k Z_i \right) \middle| Z_1 = 0 \right] \right\} (p_1 \exp(t) + 1 - p_1).
\end{aligned}$$

Repeating this inductively for the variables  $Z_2, \dots, Z_k$  yields

$$\mathbf{E}[\exp(tZ)] < \prod_i (P_i \exp(t) + 1 - P_i) = \prod_i (1 + P_i(\exp(t) - 1)).$$

Using  $1 + x < e^x$  and rearranging gives  $\mathbf{E}[\exp(tZ)] < \exp((\exp(t) - 1)b)$ . Plugging this into Eq. (19) yields Claim (a). Claim (b) can be proven analogously. □